

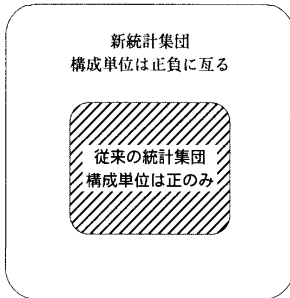
新統計集団論(2)

船津好明

(前号のあらまし)

新統計集団が従来の統計集団と異なる点は、集団の中に負の構成単位を含むことにある。負の構成単位とは、構成単位の値が負であることを指すものではなく、単位そのものが負であることを指す。従来の統計集団では構成単位がすべて正であったのに対して、新統計集団は、正負に互る構成単位からなるという点で、従来の統計集団より概念が広い。当然、新統計集団は、構成単位がすべて正となったときは、従来の統計集団と同じものになる。(図4)

図4 新統計集団の概念図



11. 期待値(続き)

前号の⑥⑩の m と⑥⑪の積率に関して期待値や確率分布を研究する。負の構成単位の存在によって、確率変数の振舞は全く目新しいものとなる。

m の期待値

⑥⑩における m の期待値を求める。⑥⑩を再掲すれば

$$m = \frac{\sum_{i=1}^n X_i U_i}{\sum_{i=1}^n U_i} \quad \text{⑥⑩} = \text{⑥⑩}$$

である。 m の期待値の算出は、 m の分母に確率変数があるため、簡単でない。2つの確率変数の比の期待値を求める問題となるので、手順を追って説明する。便宜、記号を簡素にして、 m の分母を T で、分子を W で表わす。即ち、

$$T = \sum_{i=1}^n U_i \quad \text{⑦⑩}$$

$$W = \sum_{i=1}^n X_i U_i \quad \text{⑦⑪}$$

とおくと

$$m = \frac{W}{T} \quad \text{⑦⑫}$$

となる。

まず、 T の変域や、 T が0になることがあるかどうかを検討する。各 U_i のとる値は1か-1であるので、 T は整数で、一般には

$$-n \leq T \leq n \quad \text{⑦⑬}$$

となるが、母集団の構造と標本の大きさによって、 T のとりうる値や変域は異なるものになる。そこで T の変化の態様を4つに分けて検討する。

- (1)常に $T > 0$ である場合
- (2)常に $T < 0$ である場合
- (3) $T \neq 0$ であるが $T > 0$ にも $T < 0$ にもなる場合
- (4) $T = 0$ となることがある場合

この区分において、(4)の場合を除けば、 T の実現値は有限個であるから、 m の期待値は確定する。一方母集団における μ を m に対応する特性値として選べば、 $E(m)$ と μ の比を $1 + \alpha$ とおき

$$E(m) = \mu(1 + \alpha) \tag{74}$$

と表わすことにより、 m と μ を関連づけることができる。 α は分母 T の確率変動を反映するもので、 T の確率変動がないときは $\alpha = 0$ となる。 α は状況により 0 に近いこともある、0 から大きく離れることもある。

次に上の 4 つの区分ごとに逐一検討する。

(1)常に $T > 0$ である場合

常に $T > 0$ となるには、 T の中の -1 の個数が常に n の半分を下回らなければならない。そのためには、母集団の中の -1 の個数 N_{-1} と n との関係が

$$n > 2N_{-1} \tag{75}$$

であることが必要である。すると常に $0 < T \leq n$ となる。よって m の期待値は、従来知られている 2 つの確率変数の比の期待値の公式を適用して求めることができる。即ち、 m の期待値は、 m の分子と分母の期待値を別々にとって、それらの比を作り、分母の確率変動による影響分を補正すればよい。よってまず分母は(65)により

$$E(T) = E(n\bar{U}) = n(2p - 1) \tag{76}$$

数値例 $3.5781 \ 4737 = 0.8947 \ 3684 \times 4$

となる。数値例は前号 6 頁の 8 に掲げた $N = 19$, $z = \{(x_i, u_i)\}_N$ のデータから、大きさ $n = 4$ の標本を等確率、非復元で抽出した結果である。また、分子 W は(68)により、

$$E(W) = E(n\bar{X}_U) = n\bar{x}_u \tag{77}$$

数値例 $25.0821 \ 0526 = 6.2705 \ 2632 \times 4$

となる。そして、分母の確率変動に基づく補正乗数 $1 + \alpha$ により

$$E(m) = \frac{E(W)}{E(T)}(1 + \alpha) = \frac{1 + \alpha}{N(2p - 1)} \sum_{i=1}^n x_i u_i = \mu(1 + \alpha) \tag{78}$$

$$n > 2N_{-1}, \ p \neq 0.5$$

数値例 $E(m) = 7.0599 \ 1228$

となる。(78)は(74)と同じで、 $T > 0$ の条件下での m の期待値の一般式である。数値例 7.0599 1228 は直接計算による。即ち、 m の実現値は ${}_{19}P_4 = 93024$ 通りあり、その個々に $1/93024$ をかけて総和したものである。(統計量が抽出順に関係なく、全対称型の場合は ${}_{19}C_4 = 3876$ であり。本稿の m 等は事実上抽出順に関係なく、全対称型であるので、期待値の算出においては、以下組合せの数によることとする。ただし統計量の中に $2X_1 + X_2$ のような非対称の式を含む場合は順列の数によらなければならない。)一方 $\mu = 7.0082 \ 3529$ は(45)で既に求められているから、(78)の α は

$$\alpha = \frac{E(m)}{\mu} - 1 = 0.0073 \ 7375 \tag{79}$$

となる。 $E(m)$ は μ に一致しないが、 μ との関係は(78)でわかる。

(78)における特別の場合を点検すると、 u_i のすべてが 1 の場合、 $N_{-1} = 0$ 、各 $u_i = 1$, $p = 1$ で分母の確率変動がなくなって $\alpha = 0$ となることから、この場合

$$E(m) = \frac{1+0}{N(2 \times 1 - 1)} \sum_{i=0}^N x_i \times 1 = \bar{x} = \mu \quad (80)$$

となって、 m の期待値は普通の平均に一致する。

このほか、 T と W に関する統計量の分散などは、例えば次のようになる。

$$\begin{aligned} V(T) &= E\{T - E(T)\}^2 \\ &= \frac{4(N-n)np(1-p)}{N-1} \end{aligned} \quad (81)$$

数値例 0.6648 1994

$$\begin{aligned} V(W) &= E\{W - E(W)\}^2 = \frac{(N-n)n}{N-1} \\ &\times \left\{ \frac{1}{N} \sum_{i=1}^N x_i^2 - \left(\frac{1}{N} \sum_{i=1}^N x_i u_i \right)^2 \right\} \\ &= \frac{(N-n)n}{N-1} \{ \bar{x}^2 - \bar{x}_u^2 \} \end{aligned} \quad (82)$$

数値例 40.9005 4294

$$\begin{aligned} cov(W, T) &= E\{W - E(W)\} \{T - E(T)\} \\ &= \frac{(N-n)n}{N-1} \left(\frac{1}{N} \sum_{i=1}^N x_i \right. \\ &\quad \left. - \frac{2p-1}{N} \sum_{i=1}^N x_i u_i \right) \\ &= \frac{(N-n)n}{N-1} \{ \bar{x} - (2p-1)\bar{x}_u \} \end{aligned} \quad (83)$$

数値例 4.2457 9871

次に、 $T > 0$ の場合の(74)または(78)の α を標本から求める方法を検討する。 $n=N$ や $p=1$ や $p=0$ なら分母の確率変動がなくなるから、 $\alpha=0$ となるが、そのほかの α は 0 ではなく、面倒な形となる。 α の形を具体化するには、前記の条件(1)に合った比の期待値の公式を利用しなければならない。一般理論は次のとおりである。記号を改める。

「2つの確率変数を \hat{Y}_1 、 \hat{Y}_2 とし、特に \hat{Y}_2 の変動を正で有限とする。そして各期待値を $E(\hat{Y}_1) = Y_1$ 、 $E(\hat{Y}_2) = Y_2$ 、2次の積率を $\delta_{20} = V(\hat{Y}_1)$ 、 $\delta_{11} = cov(\hat{Y}_1, \hat{Y}_2)$ 、 $\delta_{02} = V(\hat{Y}_2)$ とする。

次に $\hat{R} = \hat{Y}_1 / \hat{Y}_2$ 、 $R = Y_1 / Y_2$ とおけば \hat{R} の期待値は、2次の積率近似で

$$E(\hat{R}) \doteq R \left\{ 1 + \frac{1}{k_0^2} \left(\frac{\delta_{02}}{Y_2^2} - \frac{\delta_{11}}{Y_1 Y_2} \right) \right\} \quad (84)$$

となる。 k_0 は最適触媒値と呼ばれるもので、 $k_0 > 0.5$ である。 \hat{Y}_2 の変動範囲が正で期待値の2倍を下回るという過去の条件のもとでは、従来 $k_0 = 1$ としていたが、 \hat{Y}_2 の変動範囲の制約を緩和して、正で有限としたことによって、 k_0 は 0.5 より大きい値に置き換えられる。 $k_0 < 1$ のときは \hat{R} の R に対する偏りが、 $k_0 = 1$ の場合よりも大きくなる。 k_0 は分母の分布の歪みを結果に反映させる役割をもち、 \hat{Y}_2 の最小値と最大値の中点と Y_2 の比が現実的な値となる。よって \hat{Y}_2 の分布が対称形ならば、 $k_0 = 1$ が現実的な値となる。」(参考文献(1)、(2))

これらを前記の α に関連づけると、 $\hat{Y}_1 = W$ 、 $\hat{Y}_2 = T$ 、 $Y_1 = E(W)$ 、 $Y_2 = E(T)$ 、 $\delta_{02} = V(T)$ 、 $\delta_{11} = cov(W, T)$ 、 $\hat{R} = W/T$ 、 $R = E(W)/E(T)$ であるから、

$$\begin{aligned} \alpha &\doteq \frac{1}{k_0^2} \left(\frac{\delta_{02}}{Y_2^2} - \frac{\delta_{11}}{Y_1 Y_2} \right) \\ &= \frac{1}{k_0^2} \left\{ \frac{V(T)}{E(T)^2} - \frac{cov(W, T)}{E(W)E(T)} \right\} \\ (76), (77), (81), (83) &\text{により} \end{aligned}$$

$$\begin{aligned} &= \frac{N-n}{k_0^2(N-1)n(2p-1)} \\ &\times \left(\frac{1}{2p-1} - \frac{\sum_{i=1}^N x_i}{\sum_{i=1}^N x_i u_i} \right) \end{aligned} \quad (85)$$

数値例 $\alpha \doteq 0.0046$ 0541 ($k_0 = 1$),
0.0184 2162 ($k_0 = 0.5$)

となる。特別の場合として、 u_i のすべてが正のときは各 $u_i = 1$ 、 $p = 1$ であるから、(85)の α の右辺は 0 になって、(76)は $E(m) \doteq \mu$ ($\alpha \doteq 0$ による)となる。これは(85)を介しての結果であるが、各 $u_i = 1$ であれば、元々 $\alpha = 0$ で、正確に

$E(m) = \mu$ となる。

上例での $\alpha \approx 0.0046\ 0541$ 等は、正確には⑦の $0.0073\ 7375$ である。 $0.0046\ 0541$ 等は近似値であって、⑦の値とは一致していない。差の原因は⑤の右辺の中に3次以上の積率が入っていないことと、 k_0 のきめ方による。⑤の右辺に3次以上の項をすべてつけ加えれば、 k_0 の値にかかわらず、真の値との差はなくなる。なぜなら、⑤の元は k_0 に関係しない、 α に収斂する無限級数であるからである。なお、この種の無限級数の収斂は、必ずしも単調ではない。近似の項を1つ増やせば常に近似がよくなるかという、必ずしもそうではなく、項の増加と共に起伏を経て真の値に収斂していく形をとることがある。

— m の分布の検証例 1 —

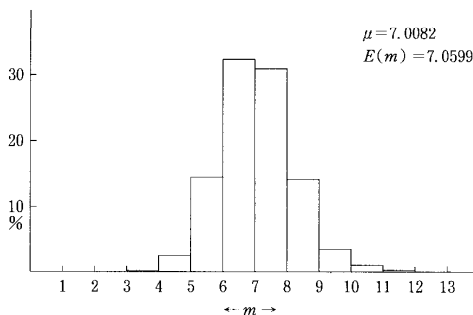
前号6頁の8で設定した検証用の大きさ $N = 19$, $u_i = -1$ の個数 $N_{-1} = 1$, $z = \{(x_i, u_i)\}_N$ の母集団から、等確率、非復元抽出した大きさ $n = 4$ の標本に基づく標本平均⑩の確率分布を追求してみる。母集団のいくつかの特性値や標本による m 以外の統計量の期待値は、前号⑤~⑥、⑫~⑬に掲げてある。 T は 4, 2 と変化する。 m に対する可能な標本は ${}_{19}C_4 = 3876$ 通りあり、その個々に対して m が計算されるから、 m も 3876 個できる。この場合の m の、刻み 1 による度数分布の表と図はそれぞれ表 1、図 5 のとおりである。度数は 7 前後に多く集まっています、形は対称形に近い。 x_i の最小値は 3.01、最大値は 11.34 であるが m は 3.01 以下、11.34 以上に互って分布している点に注意を要する。

以上が⑩の分母が常に正の場合、即ち常に $T > 0$ の場合の m の期待値や確率分布に関する説明である。

表 1 m の度数分布表

m	実度数	比率%
以上~未満		
2~3	1	0.03
3~4	12	0.31
4~5	99	2.55
5~6	562	14.50
6~7	1256	32.40
7~8	1199	30.93
8~9	548	14.14
9~10	138	3.56
10~11	45	1.16
11~12	15	0.39
12~13	1	0.03
計	3876	100.00

図 5 (表 1 に対応)



$E(m)$ の値は度数分布表からの計算ではなく、直接計算による。後現われる期待値の数値も度数分布表によらず、直接計算による。

(2)常に $T < 0$ である場合

この場合、 m の分子、分母に -1 をかけ、

$$m = \frac{-W}{-T} = \frac{\sum_{i=1}^n X_i (-U_i)}{\sum_{i=1}^n (-U_i)} \tag{86}$$

とすれば、分母は常に正になるから、(1)の理論がそのままあてはまる。よって説明は省略し(1)に準じて検証例のみを示す。

— m の分布の検証例 2 —

前号の6頁の8で設定した母集団 $N = 19$, $z = \{(x_i, u_i)\}_N$ を対比的に変更し、 u_i の符号をす

べて逆転し、次のようにする。即ち $z = \{ (3.01, -1), (4.15, -1), (4.22, -1), (4.84, -1), (5.09, -1), (5.83, 1), (6.07, -1), (6.46, -1), (6.61, -1), (6.96, -1), (7.11, -1), (7.45, -1), (7.79, -1), (8.09, -1), (8.14, -1), (8.72, -1), (9.37, -1), (9.55, -1), (11.34, -1) \}$
 $u_i = -1$ の個数 $N_{-1} = 18$ に注意を要する。これより大きさ $n=4$ の標本を等確率、非復元抽出し、 m の個々の値を計算する。 T は $-4, -2$ と変化する。 m の度数分布の表や図は検証例 1 の表 1、図 5 と同じであり、 $E(m) = 7.0599$ も変わらない。

(3) $T \neq 0$ であるが $T > 0$ にも $T < 0$ にもなる場

合

m の期待値は、その定義により

$$E(m) = \sum \frac{W}{T} \frac{1}{nC_n} \tag{87}$$

であるが、 $T > 0$ の場合と $T < 0$ の場合があるから、これらを分けて検討する。なお、 $T \neq 0$ が確保されるのは n が奇数の場合である。 n が偶数の場合、 T は常に正、常に負、正負零など色々な状況になりうる。

母集団が確定して標本の大きさが定まると、 $T > 0$ になる標本と $T < 0$ になる標本も確定するので、これらを分けて 2 つの層と考えることとする。そこで全標本を S 、そのうち $T > 0$ となる標本を S_p 、その大きさを P_o 、 $T < 0$ と

なる標本を S_n 、その大きさを N_e とする。更に $T > 0$ のときの m の条件付期待値を m_p 、 $T < 0$ のときの m の条件付期待値を m_n と表わせば、 m の期待値は m_p と m_n の加重平均となる。即ち、先ず m の S_p 内での期待値は

$$E_p(m) = \sum_{T>0} \frac{W}{T} \frac{1}{P_o} = m_p \tag{88}$$

となり、 m_p は常に $T > 0$ により、前記(1)の方法により求めることができる。

次に m の S_n 内での期待値は

$$E_n(m) = \sum_{T<0} \frac{W}{T} \frac{1}{N_e} = m_n \tag{89}$$

となり、 m_n は常に $T < 0$ により、前記(2)の方法により求めることができる。これらにより m は

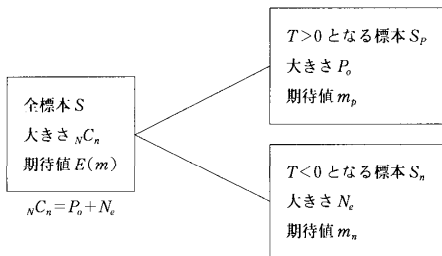
$$\begin{aligned} E(m) &= \sum \frac{W}{T} \frac{1}{nC_n} \\ &= \sum_{T>0} \frac{W}{T} \frac{1}{nC_n} + \sum_{T<0} \frac{W}{T} \frac{1}{nC_n} \\ &= \left(\sum_{T>0} \frac{W}{T} \frac{1}{P_o} \right) \frac{P_o}{nC_n} \\ &\quad + \left(\sum_{T<0} \frac{W}{T} \frac{1}{N_e} \right) \frac{N_e}{nC_n} \\ &= \frac{P_o m_p + N_e m_n}{nC_n} \end{aligned} \tag{90}$$

となる。以上により、 $T \neq 0$ で $T > 0$ にも $T < 0$ にもなる場合の m の期待値は前記(1)及び(2)の方法により求めることができる。次に検証例を示す。

— m の分布の検証例 3 —

$N = 18$ とし、 $z = \{ (x_i, u_i) \}_N$ の内容を、次のとおりとする。 $u_i = -1$ の個数 $N_{-1} = 3$ に注意を要する。 $z = \{ (3.01, 1), (4.15, 1), (4.22, 1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, 1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, 1), (7.79, 1), (8.09, 1), (8.14, -1), (8.72, 1), (9.37, 1), (9.55, 1) \}$ これを母集団として大きさ $n=5$ の標本を等確率、非復元抽出

図 6 分母 T の符号による標本の層化の構図



出し、 m の確率分布を追求してみる。可能な標本は ${}_{18}C_5=8568$ 通り、 T は5, 3, 1, -1と変化する。このうち $T>0$ となる標本は $P_0=8463$ 通りと大部分を占め、 $T<0$ となる標本は $N_0=105$ 通りと少ない。それらの度数分布の表と図は表2、図7のようになる。 m が負に互って分布している点に注意を要する。図7は図7aと図7bを併合したものであるが、総度数のほと

んどが図7aにあるため、図7と図7aは似た形状となっている。各 $E(m)$ 等の値は度数分布表によらず、直接計算による。 $E(m)=6.3522$, $E_p(m)=6.3322=m_p$, $E_s(m)=7.9627=m_n$ で、

$$\frac{8463 \times 6.3322 + 105 \times 7.9627}{8568} = 6.3522 = E(m) \quad \text{㊦}$$

表2 m の度数分布表

m	全体 ($T \neq 0$)		$T > 0$		$T < 0$	
	実度数	比率%	実度数	比率%	実度数	比率%
-4 ~ -2	9	0.11	9	0.11	0	0
-2 ~ 0	47	0.55	47	0.56	0	0
0 ~ 2	134	1.56	134	1.58	0	0
2 ~ 4	327	3.82	320	3.78	7	6.67
4 ~ 6	2440	28.48	2421	28.61	19	18.10
6 ~ 8	4789	55.89	4763	56.28	26	24.76
8 ~ 10	675	7.88	645	7.62	30	28.57
10 ~ 12	112	1.31	97	1.15	15	14.29
12 ~ 14	33	0.39	25	0.30	8	7.62
14 ~ 16	2	0.02	2	0.02	0	0
計	8568	100.00	8463	100.00	105	100.00

図7 (表2に対応)

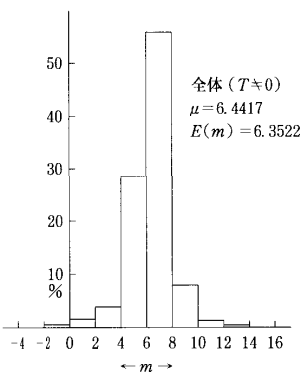


図7a

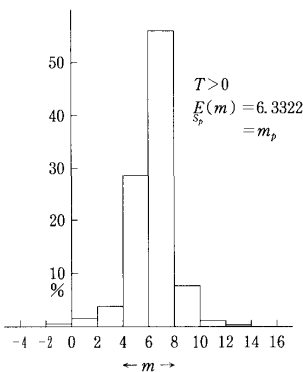
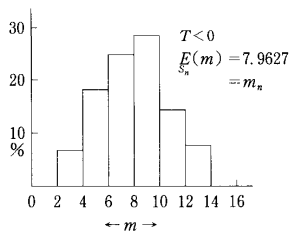


図7b



が成り立ち、式⑨が検証される。

— m の分布の検証例 4 —

$T \neq 0$ であるが $T > 0$ にも $T < 0$ にもなる場合について、もう一例を示す。 $N=18, n=5, u_i = -1$ の個数 $N_{-1}=8$ として、 $Z = \{(x_i, u_i)\}_N = \{(3.01, 1), (4.15, 1), (4.22, -1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, -1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, 1), (7.79, -1), (8.09, 1), (8.14, -1), (8.72, -1)\}$

(9.37, 1), (9.55, -1) } とおく。抽出は等確率、非復元で行われる。 T は 5, 3, 1, -1, -3, -5 と前例より大幅に変化する。 T の分散が大きくなることで、 m の分布範囲も一層広がっている。(表 3, 図 8) 可能な標本は ${}_{18}C_5 = 8568$ で前例と同じだが、 $T > 0$ となるのは 5292 通り、 $T < 0$ となるのは 3276 通りで、前例に比べ後者が大幅に増えている。この例では $E(m)$ と μ が大きく異なっている。

表 3 m の度数分布表

m	全体 ($T \neq 0$)		$T > 0$		$T < 0$	
	実度数	比率%	実度数	比率%	実度数	比率%
-8 ~ -6	2	0.02	2	0.04	0	0
-6 ~ -4	22	0.26	22	0.42	0	0
-4 ~ -2	111	1.30	111	2.10	0	0
-2 ~ 0	296	3.45	293	5.54	3	0.09
0 ~ 2	576	6.72	546	10.32	30	0.92
2 ~ 4	911	10.63	797	15.06	114	3.48
4 ~ 6	1972	23.02	1656	31.29	316	9.65
6 ~ 8	2291	26.74	1368	25.85	923	28.17
8 ~ 10	1145	13.36	342	6.46	803	24.51
10 ~ 12	639	7.46	120	2.27	519	15.84
12 ~ 14	381	4.45	31	0.59	350	10.68
14 ~ 16	166	1.94	4	0.08	162	7.95
16 ~ 18	49	0.57	0	0	49	1.50
18 ~ 20	7	0.08	0	0	7	0.21
計	8568	100.00	5292	100.00	3276	100.00

図 8 (表 3 に対応)

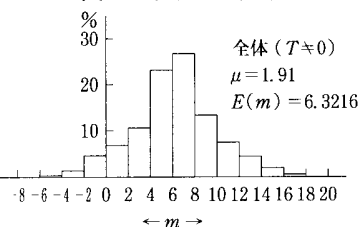


図 8 a

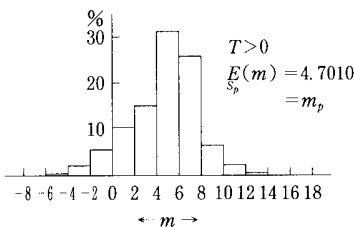
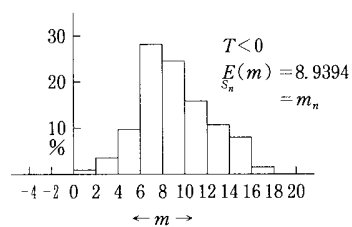


図 8 b



(4) $T=0$ となることがある場合

$T=0$ は、 n が偶数の場合に起こることがある。 n が奇数のときには起きない。 $T=0$ になると、これと共に $T>0$ または $T<0$ の一方または両方が起きる。この場合の m の期待値について検討する。

$$m = \frac{W}{T} \tag{92} = \text{⑦②}$$

において、 m の期待値は、 m のすべての実現値にその実現確率を掛けて総和したものであるから、有限の標本集団においては、 $T=0$ が1個でも発生すると期待値はなくなる。よって本稿における抽出条件では、 $T=0$ が発生すると m の期待値は存在しない、というのが結論である。

しかし、これで検討を終えるのではなく、 T のうち0となるものを除いた標本集団について、期待値に代わるものを考えることとする。

すべての標本からなる集団 S の大きさは ${}_N C_n$ であるが、この中の $T=0$ となるような標本の個数を t_0 とおき、これを m に対する無効標本と呼ぶ。これにより ${}_N C_n - t_0$ は m に対する有効な標本からなる集団の大きさとなる。この集団を S_v と表わす。 S_v は標本の全集団 S に対する代替集団と位置づける。よって有効標本の個々が抽出される確率を、無効標本を除くという条件付確率に置きかえ、 $1/({}_N C_n - t_0)$ とする。併

せて m' (ダッシュ) をつけて記し、本来の m と区別する。これらにより、集団 S_v における m' の期待値は

$$E(m') = \sum_{T \neq 0} \frac{W}{T} \frac{1}{{}_N C_n - t_0}, \quad t_0 > 0 \tag{93}$$

となる。標本の個数は有限個であるから、この値は必ず求まる。この場合、 T は正のみ、または負のみ、または正負両方に互って変動するのが一般で、よって⑨③は前記(1)、(2)、(3)の方法によって求めることができる。

— m の分布の検証例5 —

$N=19$ として、 $z = \{(x_i, u_i)\}_N$ の内容を、

表4 m' の度数分布表

m'	実度数	比率%
	以上~未満	
2~3	7	0.19
3~4	40	1.07
4~5	157	4.20
5~6	577	15.43
6~7	1116	29.84
7~8	1057	28.26
8~9	538	14.39
9~10	170	4.55
10~11	60	1.60
11~12	17	0.45
12~13	1	0.03
計	3740	100.00

図9 標本の有効・無効の構図

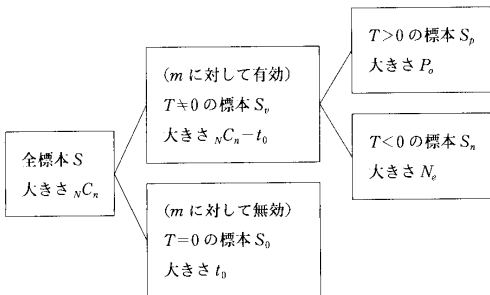
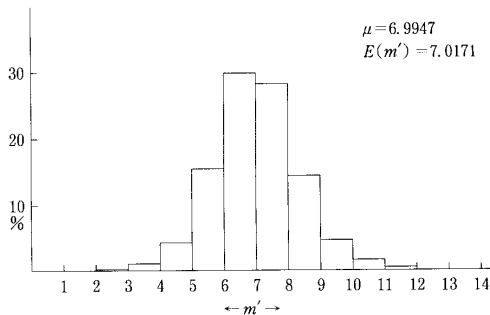


図10 (表4に対応)



前例と異なり次の通りとする。 $u_i = -1$ の個数 $N_{-1} = 2$ に注意を要する。 $z = \{(3.01, 1), (4.15, 1), (4.22, 1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, 1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, 1), (7.79, 1), (8.09, 1), (8.14, 1), (8.72, 1), (9.37, 1), (9.55, 1), (11.34, 1)\}$ これを母集団として、大きさ $n = 4$ の標本を等確率、非復元抽出する。この場合、 T は4, 2, 0と変化する。 $T = 0$ となる標本は136通りあり、これを無効とする。有効標本は、全標本 ${}_{19}C_4 = 3876$ 通りから136通りを引いた3740通りとなる。これに対しては常に $T > 0$ である。 $E(m')$ の値は直接計算によれば $E(m') = 7.0171$ となる。 m' の度数分布の表と図は、表4、図10の通りである。この例では $E(m')$ と μ の差はあ

まりない。

— m の分布の検証例6 —

$N = 19$ として、 $z = \{(x_i, u_i)\}_N$ の内容を、前例と異なり次の通りとする。 $u_i = -1$ の個数 $N_{-1} = 3$ に注意を要する。 $z = \{(3.01, 1), (4.15, 1), (4.22, 1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, 1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, 1), (7.79, 1), (8.09, 1), (8.14, -1), (8.72, 1), (9.37, 1), (9.55, 1), (11.34, 1)\}$ これを母集団として、大きさ $n = 4$ の標本を等確率、非復元抽出する。この場合、 T は4, 2, 0, -2と変化する。 $T < 0$ が生じている点が前例と異なる。 $T = 0$ となる標本は360通りあり、これらを無効とする。有効標本は、全標本 ${}_{19}C_4 = 3876$ 通りから360通りを引いた3516通りと

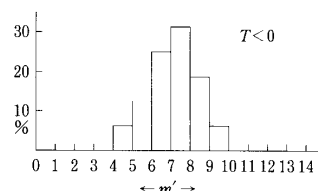
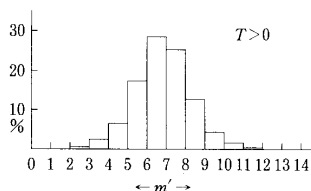
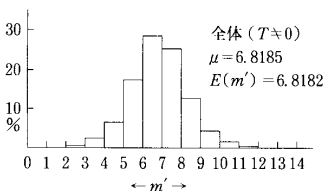
表5 m' の度数分布表

m'	全体 ($T \neq 0$)		$T > 0$		$T < 0$	
	実度数	比率%	実度数	比率%	実度数	比率%
1~2	3	0.09	3	0.09	0	0
2~3	21	0.60	21	0.60	0	0
3~4	86	2.45	86	2.46	0	0
4~5	230	6.54	229	6.54	1	6.25
5~6	612	17.41	610	17.43	2	12.50
6~7	1000	28.44	996	28.46	4	25.00
7~8	888	25.26	883	25.23	5	31.25
8~9	453	12.88	450	12.86	3	18.75
9~10	152	4.32	151	4.31	1	6.25
10~11	55	1.56	55	1.57	0	0
11~12	15	0.43	15	0.43	0	0
12~13	1	0.03	1	0.03	0	0
計	3516	100.00	3500	100.00	16	100.0

図11 (表5に対応)

図11a

図11b



なる。このうち $T > 0$ となるのは 3500 通りと大部分を占め、 $T < 0$ となるのは 16 通りと僅かである。 m' の度数分布の表と図は、表 5、図 11 の通りである。図 11b は実度数が少ないから、全体への影響は少ない。この例では $E(m')$ と μ の差はあまりない。

— m の分布の検証例 7 —

$N=19$ として、 $z = \{(x_i, u_i)\}_N$ の内容を、前例と異なり次のとおりとする。 $u_i = -1$ の個数 $N_{-1}=9$ に注意を要する。 $z = \{(3.01, 1), (4.15, 1), (4.22, -1), (4.84, 1), (5.09, 1), (5.83, -1), (6.07, 1), (6.46, -1), (6.61, 1), (6.96, 1), (7.11, -1), (7.45, -1), (7.79, -1), (8.09, 1), (8.14, -1), (8.72, -1), (9.37, 1), (9.55, -1), (11.34, 1)\}$ これを母集団として、

大きさ 4 の標本を等確率、非復元抽出する。この場合、 T は 4, 2, 0, -2, -4 と前例より大幅な変化をする。 $T=0$ となる標本は 1620 通りあり、これを無効とする。有効標本は、全標本 ${}_{19}C_4=3876$ 通りから 1620 通りを引いた 2256 通りとなる。このうち $T > 0$ となるのは 1290 通り、 $T < 0$ となるのは 966 通りである。 m' の度数分布の表と図は、表 6、図 12 の通りである。この例では $E(m')$ と μ は大きく異なっている。

μ が存在しない場合

これまでの母集団では μ が存在した。そして m の期待値は④等によって μ と関連づけることができた。ところが μ が存在しないときは④等は成立しない。 m の期待値は単に m の

表 6 m' の度数分布表

m'	全体 ($T \neq 0$)		$T > 0$		$T < 0$	
	実度数	比率%	実度数	比率%	実度数	比率%
以上～未満						
1～2	7	0.31	7	0.54	0	0
2～3	44	1.95	42	3.26	2	0.21
3～4	114	5.05	102	7.91	12	1.24
4～5	214	9.49	174	13.49	40	4.14
5～6	345	15.29	252	19.53	93	9.63
6～7	460	20.39	273	21.16	187	19.36
7～8	476	21.10	218	16.90	258	26.71
8～9	320	14.18	125	9.69	195	20.19
9～10	190	8.42	64	4.96	126	13.04
10～11	72	3.19	26	2.01	46	4.76
11～12	13	0.58	6	0.47	7	0.72
12～13	1	0.04	1	0.08	0	0
計	2256	100.00	1290	100.00	966	100.00

図 12 (表 6 に対応)

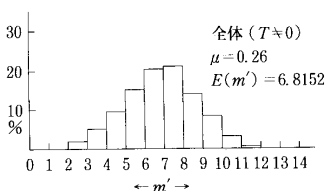


図 12a

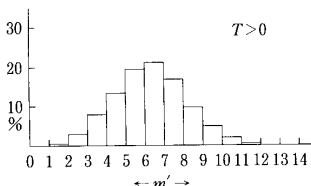


図 12b

